The impact of targeted Rheumatoid Arthritis pharmacological treatment on mental health: A systematic review and network meta-analysis.

Faith Matcham, PhD[1,], James Galloway, PhD[2,3], Matthew Hotopf, PhD[1], Emmert Roberts, BMBCh[1,4], Ian C Scott, PhD[5], Sophia Steer, PhD[3], Sam Norton, PhD[3,6].

1. Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK.
2. Academic Rheumatology, Faculty of Life Sciences and Medicine, King's College London, UK
3. Rheumatology Department, King's College Hospital NHS Foundation Trust, London, UK.
4. National Addiction Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK.
5. Primary Care and Health Sciences, Keele University, UK.
6. Health Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London.

Corresponding author:

Dr Faith Matcham

Department of Psychological Medicine, Weston Education Centre, 10 Cutcombe road, London SE5 9RJ.

Telephone: 02078480868

Email: Faith.Matcham@kcl.ac.uk

Word Count: 3,671

ABSTRACT

Rheumatoid Arthritis (RA) pharmacotherapy may impact mental health (MH) outcomes by improving pain and stiffness; and potentially via targeting inflammatory processes common to RA and depression. The objectives of this review were to i) ascertain the frequency of MH assessment in RA pharmacotherapy trials; ii) quantify the efficacy of RA pharmacotherapy efficacy on MH outcomes; iii) explore the clinical and demographic factors related to MH outcomes.

CENTRAL, PsychINFO, Web of Science, Medline, Embase and CINAHL were systematically searched from inception to March 2017 for randomised trials of disease-modifying anti-rheumatic drugs (DMARDs) in adult RA patients. The primary outcome was MH; self-reported physical health was extracted as a secondary outcome. Pairwise meta-analysis (PMA) created pooled effect sizes and 95%CIs for comparisons of all treatments versus comparators (active or placebo). Network meta-analysis (NMA) provided effect size estimates of targeted biologic DMARDs (bDMARDs) versus conventional synthetic DMARDs (csDMARDs) using indirect comparisons of different treatment modalities.

71 eligible studies were identified. 57 studies were included in the PMA, representing 23,535 patients. bDMARDs showed small effects on MH (standardised mean difference (SMD) versus csDMARDs = 0.19 to 0.30), and moderate effects on self-reported physical health (SMD versus csDMARDs = 0.46 to 0.50), with NMA determining no significant differences in effectiveness between bDMARD mode of action on either outcome.

Effective pharmacotherapy alone is unlikely to substantially improve MH outcomes for most RA patients. Integrated MH care provided within routine clinical practice is essential to optimise mental and physical health outcomes.

DECLARATION OF INTERESTS

None

Keywords: anti-tnf, DMARDs (biologic), rheumatoid arthritis, psychology, treatment

INTRODUCTION

Rheumatoid Arthritis (RA) is an autoimmune disease with a prevalence of 0.5-1.0% in adults. [1] RA causes swelling and pain of the joints (mainly hands, wrists and feet) reducing functional ability, which can substantially impact both physical and mental quality-of-life (QoL; [2]). Mental health (MH) disorders are highly prevalent; approximately 17% of RA patients have depressive disorder according to diagnostic interview [3] and 25.1% of rheumatology outpatients screen positive for anxiety disorder. [4] These estimates are substantially higher than for the general population, where depression prevalence estimates are typically around 5%. [5] Poor MH is associated with numerous deleterious outcomes in RA; increased risk of mortality, [6] work disability, [7] worsened disease activity and physical function, [8–10] higher pain and [11] fatigue. [12]

There is increasing evidence suggesting common inflammatory pathways between RA and depression. Specifically, inflammatory cytokines including tumour necrosis factor alpha (TNF-α) and interleukin-6 (IL-6) can be elevated in people with depressive disorder [13] and recent evidence suggests that therapies used in RA targeting TNF-α inhibitors may improve MH outcomes in depressed patients with high levels of inflammation, [14] and with chronic physical illness. [15]

RA management has evolved in the last 25 years, with earlier diagnosis, and earlier, more aggressive treatment. [16] The “treat to target” framework emphasises the desired goal of reaching a state of remission, switching medications until this target has been achieved. [17,18] Initial treatment typically involves conventional synthetic disease-modifying anti-rheumatic drugs (csDMARDs), usually methotrexate. In the UK, more expensive targeted biologic DMARDS (bDMARDs) are reserved for those with insufficient response to two csDMARDs. [19] For the purposes of this review, we use the term bDMARDs to encompass both targeted biologic and Janis kinase inhibitor (JAK) treatments. Whilst there has been evident improvements in radiographic outcomes and inflammation, impact on physical function

and QoL is less pronounced. [20,21] The limited impact on QoL is worrying given that psychosocial wellbeing and social function are of key importance to patients. [15]

As low mood is highly prevalent in RA, [3] and psychosocial wellbeing is important to patients, [22] it might be expected that MH is commonly assessed as an outcome in RA clinical trials. However, a 2009 systematic review found that MH outcomes were reported in 4% of RA clinical studies, [23] increasing to 22% with a broader conceptualisation of mood including MH components of QoL using questionnaires such as the Medical Outcome Survey 36-item Short Form (SF36; [24]).

The aim of this study was to systematically review the evidence around the efficacy of pharmacotherapy on improving MH outcomes in RA. The objectives were to: 1) identify the frequency with which MH outcomes are measured and reported in RA pharmacotherapy trials; 2) quantify the impact of bDMARDs on MH outcomes, comparing against self-reported physical health; and 3) investigate factors that may moderate RA pharmacological treatment efficacy for MH outcomes, such as treatment mode of action, patient demographic, and clinical characteristics.

METHODS

Identification of trials

A protocol and data extraction form were developed in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA; [25]) statement (appendix 1). We searched the Cochrane Central Register of Controlled Trials (CENTRAL), PsychINFO, Medline, Embase, Web of Science and CINAHL from inception to March 2017. Search terms are available in the protocol, provided in appendix 2. We also screened reference lists of reviews and ClinicalTrials.gov for trials still in progress. Titles were screened for relevance, followed by abstracts and full-texts to assess eligibility for inclusion. This screening procedure was conducted by reviewer FM, with reviewer ER following the same procedure for 10% (460/4604) of identified articles.

Selection criteria

Types of patient

Studies reporting data from adult patients aged >18 years with RA were included. Studies spanning several disease groups were only eligible if results from RA patients were reported separately.

Study design and treatment types

Randomised controlled trials (RCTs) of bDMARD pharmacological treatments for managing RA, including drugs in use in clinical practice at the time of study and new drugs under investigation, were eligible. Generic pain relief medication or alternative and complementary therapies such as acupuncture or collagen were excluded. Trials including active comparators (bDMARD vs. bDMARD), placebo control groups (bDMARD vs. placebo) or usual care control groups (bDMARD vs. csDMARD) were included, as were multi-arm trials (bDMARD vs. bDMARD vs. csDMARD). For cross-over trials, data were extracted from the first period only, to avoid potential carryover effects. Pragmatic trials, with patients shifting between treatment modalities and dosages according to treatment response were included in a narrative synthesis.

Outcomes

Our primary outcome of interest was MH, including both traditional depression and anxiety questionnaires and generic measures of QoL that include MH subscales. Data from these questionnaires were included if they were reported from MH subscales separately from overall quality-of-life or disability scores.

Based on previous systematic review evidence, [23] we anticipated that the SF36 would be the most commonly-used questionnaire. If data were reported from more than one MH questionnaire, data from the SF36 were prioritised for inclusion in meta-analysis to reduce heterogeneity and aid interpretation. The SF36 has eight domains assessing various aspects of mental and physical well-being: physical function (PF); role physical (RP); global health

(GH); bodily pain (BP); vitality (V); social function (SF); role emotional (RE); and mental health (MH). [26] These domains can be combined to form two higher-order summary scores: Physical Component Summary (PCS); and Mental Component Summary (MCS). The PCS is formed by positively weighting the physical domains (PF, RP, GH, BP) and negatively weighting the mental domains (V, SF, RE, MH) and the MCS is calculated by positively weighting the mental domains and negatively weighting the physical domains. The PCS and MCS summary scores are inter-related, [27] yet provide an indicator of the impact of treatment on physical outcomes in comparison to mental outcomes, with higher scores indicating improved mental/physical QoL. PCS scores were considered secondary outcome data, to allow comparison between mental QoL and physical QoL outcomes following RA treatment.

Data extraction

Data were extracted from all eligible papers (N=71) by two reviewers (FM and ER) independently, to minimise human error in reporting results (appendix 3). In the case of incomplete reporting of data, we searched ClinicalTrials.gov, accessed company-specific registries, contacted authors directly, and made data requests to funding bodies as necessary.

Risk of Bias

A key assessment of the quality of the information provided by a trial is the potential for bias in the treatment effect estimate. Risk of bias of included trials was assessed by 2 reviewers (FM/ER) using the Cochrane tool. [28] This assessed random sequence generation, allocation concealment, participant, personnel and outcome assessor blinding, completeness of outcome data, and selective reporting. Where necessary, this data was obtained from “parent” primary outcome papers, where more detailed methodological information is included.

The quality of each outcome was assessed using the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) system. A rating of high, moderate, low or very low was given to each outcome (MCS and PCS), based on assessment of risk of bias, inconsistency (between estimated effect sizes across studies and estimated I^2 heterogeneity),

indirectness (applicability of study to the review aim), imprecision, and risk of publication bias (appendix 4).

Statistical methods

Standardised mean difference (SMD) effect sizes were calculated for each comparison using group means and standard deviations (SDs). The SMD indicates the size of the treatment effect relative to the observed variability in the outcome and can be interpreted as the between group difference in SD units; where an SMD of .5 indicates half a SD difference. A rule of thumb is that SMDs of .2, .5 and .8 are interpreted as small, medium, and large effects, respectively [29]. Where multiple doses of the same drug were tested, the most commonly used dosage, or dosage most reflecting clinical practice was included in the pooled meta-analysis. Where dose-finding studies of new drugs used a range of doses, the mean scores across dosages was taken. Endpoint means were prioritised, however mean change scores were included where endpoint scores were unavailable. If no mean scores *or* SDs were available after accessing ClinicalTrials.gov, or contacting authors and funding bodies, effect sizes were calculated using any available statistical estimates including t-scores, 95% confidence intervals, and p-values. [30] Missing SD data were imputed by calculating the mean SD from data available from other studies using the same outcome, drug and dosage at the same time-point.

The analysis involved random-effects pairwise meta-analysis (PMA), due to expected heterogeneity, including all studies regardless of comparator using Stata v14. Subgroup analyses compared active treatment separately with (no treatment) placebo and with csDMARD controls. Statistical heterogeneity in the between study treatment effects was assessed using I^2 , with scores of 25%, 50% and 75% representing low, moderate and high heterogeneity respectively. [31] The pooled treatment effect estimated may not be trustworthy when heterogeneity is high. Additionally, meta-regression was used to investigate between study differences in design and patient characteristics that might account for variability in between study treatment effects. Study sample size, age, proportion female, disease duration,

baseline mood, baseline disease activity, follow-up time in weeks, rheumatoid factor (RF) status, recruitment year, and availability of data were entered as a bivariate exploration in studies of bDMARDs vs csDMARDs. A significant difference between analyses was established when confidence intervals did not overlap.

Studies examining bDMARDs vs. csDMARDs were used in network meta-analysis (NMA) of targeted therapies by mode of action. NMA is an extension of traditional PMA to multiple treatment comparisons, which allows indirect comparisons to be made between different treatment types. [32] For example, if etanercept and abatacept have both been compared directly to MTX in different trials, the relative effectiveness of etanercept versus abatacept can be estimated indirectly. This method also has the benefit of combining direct and indirect comparisons to provide a more precise (i.e. smaller standard errors) estimate of effect size. [32]

Since the NMA grouped treatment by mode of action, it was necessary to exclude studies comparing bDMARDs with the same mode of action without a csDMARD or placebo control arm. Typically, such studies concerned a bDMARD biosimilar. Effect sizes were presented as pooled SMD and 95% CIs. Direct and indirect estimates of effect size were compared for bDMARD subcategories where direct comparisons were available, and comparison-adjusted funnel plots were created to indicate differences in effect sizes between small and large studies. Targeted treatments were ranked based on the estimated probability of each targeted treatment being most effective for MCS and PCS outcomes, which was estimated using surface under the cumulative ranking curve (SUCRA). SUCRAs combine the estimated probabilities (derived from the NMA) that each treatment is the first best, second best, and so on for all possible ranks (provided in web appendix 4). Higher SUCRA values indicate greater likelihood of a given treatment being the most efficacious, such that where the SUCRA is one the treatment is certain to be the best, and where it is zero is certain to be the worst.

RESULTS

Search results and included participants

A total of 71 studies, involving 34,796 participants, were identified (figure 1/table1). Full references for these studies are provided in appendix 6. The mean age of patients ranged from 47 to 57.5 years, 78.6% female, and the mean disease duration ranged between 0.1 and 12.3 years. The mean baseline MCS scores was 42.2 and the mean baseline DAS-28 was 6.2. The studies considered 16 bDMARDs: anti-TNFs (adalimumab, certolizumab, etanercept, golimumab, infliximab); B-cell inhibitors (rituximab, SBI-087); T-cell inhibitors (abatacept); anti-IL6 (clazakizumab; sarilumab; sirukumab; tocilizumab) and Janus Kinase inhibitors (baricitinib, decernotinib, fostamitinib, tofacitinib).

Objective 1: The frequency of MH outcome measurement

Of the 71 eligible studies, with evidence of mood having been measured in either an abstract, methods, or as a list of outcomes on ClinicalTrials.gov, only 36 (50.7%) reported MH data in either publications, supplementary material, or open online data summary reports. Attempts were made to contact authors and funders of 32 of the remaining 35 studies with insufficient information available (3 papers did not have contact information or funding information available); only 12 (36.4%) of these contact attempts resulted in receipt of the necessary data. Of the remaining 23 where no data were available, imputation of the missing information (e.g. SD of the outcome) was possible for 12 studies (allowing inclusion in the meta-analysis), 4 reported some data which were added to the narrative synthesis (appendix 7), and 7 were not able to be included in any outcome assessment. A total of 57 papers were included in the PMA and 54 in the NMA. The three studies omitted from the NMA were head-to-head trials of targeted therapies in the same class.

Table 1. Study characteristics.

Study ID	Analysis Inclusion	Interventions	Year	Patient N	Female N (%)	Mean Age, (SD)	Mean disease duration (SD)	Follow-up (weeks)	Missing data	Mood measurement	Baseline mood, mean (SD)
ADACTA	Meta-analysis	Tocilizumab (8mg/kg q4w) vs adalimumab (40mg eow)	2010-2011	325	262 (80.6)	53.4 (12.7)	6.8	24	LOCF	SF36	-
AIM	Meta-analysis	Placebo (+MTX) vs abatacept (10mg/kg q4w)		652	516 (79.1)	51.0 (12.7)	8.7 (7.2)	52	LOCF	SF36	41.3 (11.3)
Alemao 2014	Meta-analysis	MTX vs clazakizumab (25-200mg q4w) MTX vs adalimumab (40mg eow)	-	418	-	-	-	24	-	SF36	-
AMPLE	Meta-analysis	Abatacept (125mg/wk) vs adalimumab (40mg eow)	-	646	529 (81.9)	51.2 (12.7)	1.8 (1.4)	104	Excluded	SF36	43.5 (11.5)
APPEAL	Meta-analysis	DMARD+MTX vs etanercept (50mg/wk)	2007-2009	300	271 (90.3)	48.5 (11.7)	6.2 (7.9)	16	LOCF	SF36	42.7
ATTAIN	Meta-analysis	Placebo (+MTX) vs abatacept (10mg/kg q4w)	2002-2004	393	305 (78.0)	53.1 (11.9)	11.8 (8.7)	24	LOCF	SF36	42.1 (12.2)
ATTEST	Meta-analysis	Placebo (+MTX) vs abatacept (10mg/kg) Placebo vs infliximab (3mg/kg)	-	431	362 (84.0)	49.2 (12.0)		28	LOCF	SF36	-
ATTRACT	Narrative Synthesis	Placebo (+MTX) vs infliximab (3mg/kg q8w-10mg/kg q4w	1997-1998	428	332 (78.0)	54	10.6 (8.4)	102	-	SF36	Median = 48.1

AVERT	Meta-analysis	MTX vs abatacept (125mg/wk)	-	511	273 (77.8)	47.0 (12.6)	0.6 (0.5)	24, 52	Imputation	SF36	41.3 (11.2)
BEST	Narrative Synthesis	Sequential monotherapy vs step-up combination therapy vs initial combination therapy + prednisolone vs initial combination therapy + infliximab	2000-2002	508	343 (67.5)	54.4 (13.8)	Median = 0.5	12, 24, 52, 104	ITT	SF36	47.3
Burmester 2013	Excluded	Placebo (+MTX) vs mavrilimumab (100mg eow)	-	139	-	-	-	4, 12	-	SF36	-
CERTAIN	Meta-analysis	Placebo (+MTX) vs certolizumab (200mg)	2008-2010	194	156 (80.4)	53.8 (12.2)	4.6 (3.4)	24	LOCF	SF36	43.2 (10.7)
Choy 2012	Meta-analysis	Placebo (+MTX) vs certolizumab (400mg)	2002-2004	247	171 (69.2)	54.3 (12.0)	9.7 (7.7)	24	LOCF	SF36	45.7 (12.3)
COMET	Meta-analysis	MTX (7.5mg-50mg/wk) vs etanercept (50mg/wk)	2004-2006	542	387 (73.0)	51.4 (13.8)	0.8 (0.5)	52	LOCF	SF36, HADS	SF36: 42.2 (12.0) HADS (dep): 6.8 (4.1) HADS (anx): 7.5 (4.4)
CONCERTO	Narrative Synthesis	MTX (2.5mg/5mg/10mg/20 mg) vs adalimumab (40mg eow)	2010-2012	395	300 (75.9)	51.9 (13.4)	0.3 (0.4)	26	LOCF	SF36	-

Damjanov 2016	Meta-analysis	SBI-087 (200mg) + MTX vs placebo (+MTX)	-	209	164 (78.5)	54.7 (12.2)	8.5 (7.8)	16, 24	LOCF	SF36	-
DANCER	Meta-analysis	Placebo (+MTX) vs rituximab (2x500mg) Placebo (+MTX) vs rituximab (2x1000mg)	-	367	287 (78.2)	51.4 (11.6)	10.7 (8.2)	24	Excluded	SF36	41.4 (12.0)
Durez 2004	Meta-analysis	Infliximab (3mg/kg) (+MTX) vs methylprednisolone (1g) (+MTX)	-	27	23 (85.2)	Median =42.0	Median =11.0	14	-	SF36	48.5
Emery 2006	Meta-analysis	Placebo (+MTX) vs abatacept (10mg/kg) Placebo (+MTX) vs abatacept (2mg/kg)	-	339	230 (68.0)	55	9.4 (8.7)	52	LOCF	SF36	43.8 (12.7)
FAST4WARD	Meta-analysis	Placebo vs certolizumab (400mg)	2003-2004	220	184 (83.6)	53.8 (12.2)	9.6 (8.9)	24	Imputation	SF36	44.7 (11.5)
FUNCTION	Meta-analysis	Placebo (+MTX) vs tocilizumab (4mg/kg)+MTX Placebo (+MTX) vs tocilizumab (8mg/kg)+MTX Placebo (+MTX) vs tocilizumab (8mg/kg)	-	1157	904 (78.1)	50.1 (13.5)	0.5 (0.5)	24, 52	Excluded	SF36	-
Genovese 2004	Excluded	Etanercept (25mg biw) (+MTX) vs etanercept (25mg biw) + anakinra (100mg qd) (+MTX) Etanercept (25mg biw) (+MTX) vs etanercept (25mg qw) + anakinra (100mg qd) (+MTX)	-	242	187 (77.3)	54.6 (12.8)	9.9 (9.8)	24	ITT	SF36	46.4 (11.7)

GO-FORWARD	Meta-analysis	Placebo (+MTX) vs golimumab (100mg) Placebo (+MTX) vs golimumab (50mg) Placebo (+MTX) vs golimumab (100mg)+MTX	2005-2007	444	358 (80.6)	50.4 (11.3)	8.3 (8.0)	24	ITT	SF36	43.8 (11.0)
GO-FURTHER	Meta-analysis	Placebo (+MTX) vs golimumab (2mg/kg)	2009-2011	592	483 (81.6)	51.8 (12.1)	7.0 (7.1)	24	LOCF	SF36	37.6 (11.3)
HERA	Meta-analysis	HD203 (25mg biw) vs etanercept (25mg biw) (+MTX)	2010-2012	294	202 (68.7)	51.2 (12.2)	7.7 (7.4)	24, 48	LOCF	SF36	39.8 (11.6)
HIKARI	Meta-analysis	Placebo (+MTX) vs certolizumab (200mg)	2008-2010	230	171 (74.3)	55.7 (10.0)	5.6 (4.2)	12, 24	LOCF	SF36	44.8 (12.9)
HIT HARD	Meta-analysis	Placebo (+MTX) vs adalimumab (40mg eow)	2007-2010	172	118 (68.6)	49.9 (13.2)	0.1 (0.7)	24, 48	MI	SF36	46.0 (10.1)
IMAGE	Meta-analysis	Placebo (+MTX) vs rituximab (2x500mg) Placebo (+MTX) vs rituximab (2x1000mg)	2006-2007	748	607 (81.1)	48.0 (13.1)	0.9 (1.2)	52	LOCF	SF36	36.7 (12.2)
J-RAPID	Meta-analysis	Placebo (+MTX) vs certolizumab (400mg) Placebo (+MTX) vs certolizumab (200mg) Placebo (+MTX) vs certolizumab (100mg)	2008-2010	316	262 (82.9)	53.1 (10.9)	5.9 (4.1)	24	LOCF	SF36	46.6 (11.7)
Keystone 2004	Narrative Synthesis	Placebo (+MTX) vs adalimumab (40mg biw) Placebo (+MTX) vs adalimumab (20mg biw)	-	619	464 (75.0)	56.5 (12.0)	11.0 (9.1)	12, 24, 52	ITT	SF36	-

Kim 2013	Excluded	Placebo (+MTX) vs infliximab (3mg/kg)	2005-2006	143	128 (89.5)	50.4 (10.8)	Median =8.6	30	ITT	SF36	-
Kremer 2003	Narrative Synthesis	Placebo (+MTX) vs abatacept (2mg/kg q4w) Placebo (+MTX) vs abatacept (10mg/kg q4w)	2000-2001	339	231 (68.1)	55	9.4 (8.7)	24	LOCF	SF36	43.2 (10.8)
Kremer 2014	Excluded	Placebo (+MTX) vs mavrilimumab (30-100mg eow)	-	326	282 (86.5)	51.8 (11.1)	-	12, 24	ITT	SF36	-
Li 2016	Meta-analysis	Placebo (+MTX) vs golimumab (50mg q4wks)	-	264	214 (81.1)	47.2 (11.8)	7.8 (7.2)	24	LOCF	SF36	39.6 (11.1)
Machado 2014	Meta-analysis	DMARD+MTX (7.5mg-25mg/wk) vs etanercept (50mg/wk)+MTX (7.5mg-25mg/wk)	2009-2012	423	376 (88.8)	48.5 (11.7)	8.5 (7.3)	24	LOCF	SF36	SF36: 40.0 (10.7)
Manders 2015	Excluded	Anti-TNF vs abatacept (10mg/kg q4w) Anti-TNF vs rituximab (1000mg eow)	2009-2012	144	104 (74.8)	56.3 (11.2)	Median = 6.3	24, 52	ITT	SF36	-
Mathias 2000	Meta-analysis	Placebo vs etanercept (25mg biw) Placebo vs etanercept (10mg biw)	-	234	182 (77.9)	52.3	12	2, 12, 24	LOCF	SF36	41.7
Mease 2012	Meta-analysis	Placebo (+MTX) vs clazakizumab (80-320mg)	2008-2009	127	-	52.5 (11.3)	7.0 (6.0)	16	ITT	SF36	34.5 (11.9)
MUSICA	Narrative Synthesis	MTX (7.5mg/wk)+adalimumab (40mg eow) vs MTX	-	309	-	54.8	-	24	LOCF	SF36	-

		(20mg/wk)+adalimumab (40mg eow)									
OPERA	Meta-analysis	Placebo (+MTX) vs adalimumab (40mg eow)	2007-2009	180	119 (66.1)	55.2	0.2	52	LOCF	SF36	46.9
OPTION	Meta-analysis	Placebo (+MTX) vs tocilizumab (4mg/kg)	2005-2006	623	510 (81.9)	50.9 (12.2)	7.6 (7.3)	24	Excluded	SF36	40.0 (11.1)
		Placebo (+MTX) vs tocilizumab (8mg/kg)									
ORAL	Meta-analysis	Placebo (+MTX) vs tofacitinib (5/10mg bid)	2009-2011	399	335 (84.0)	55.0 (11.4)	12.3	24	ITT	SF36	42.5 (12.9)
ORAL-SCAN	Meta-analysis	Placebo (+MTX) vs tofacitinib (5/10mg bid)	2009-2011	797	679 (85.2)	52.8 (11.6)	9.1	4, 12, 24, 52	Excluded	SF36	42.1 (11.6)
ORAL-STANDARD	Meta-analysis	Placebo (+MTX) vs tofacitinib (5/10mg bid)	2009-2011	717	586 (81.7)	53.2 (12.6)	7.8	24	Excluded	SF36	41.0 (11.3)
		Placebo (+MTX) vs adalimumab (40mg eow)									
ORAL-START	Meta-analysis	MTX (10mg-20mg/wk) vs tofacitinib (5/10mg bid)	2010-2013	956	758 (79.3)	49.5	3	52, 104	Excluded	SF36	-
ORBIT	Meta-analysis	MTX + rituximab (2x500mg) vs MTX + anti-TNF (adalimumab 40mg eow or etanercept 50mg pw)	2009-2013	295	213 (72.2)	57.0 (10.0)	7.4 (7.3)	52	ITT	HADS	-
PLANETRA	Meta-analysis	Infliximab (3mg/kg) vs biosimilar (CT-P13 3mg/kg)	-	506	501 (82.7)	Median = 50 Range 18-75	-	14, 30, 54	LOCF	SF36	37.6 (10.9)

PREMIER	Meta-analysis	MTX (20mg/wk) vs adalimumab (40mg eow) vs adalimumb (40mg eow) + MTX (20mg/wk)	-	799	595 (74.5)	52.0 (13.5)	0.7 (0.8)	12, 26, 52, 76, 104	-	SF36	43.4 (12.3)
PRIZE	Meta-analysis	Placebo (+MTX) vs MTX (10-25mg/wk) vs etanercept (25mg) +MTX	2009-2012	193	125 (64.8)	49.4 (14.4)	0.3 (0.2)	39	LOCF	SF36	43.5 (10.8)
RA-BEACON	Meta-analysis	Placebo (+DMARD) vs MTX + baricitinib (2mg or 4mg daily)	2013-2014	527	431 (81.8)	55.7 (11.0)	14.0 (9.0)	24	ITT	SF36	-
RA-BEAM	Meta-analysis	Placebo (+MTX) vs baricitinib (4mg QD) vs adalimumab (40mg q2w)	2012-2014	1305	1008 (77.2)	53.3 (5.3)	10.0 (9.0)	24	LOCF	SF36	-
RADIATE	Meta-analysis	Placebo (+MTX) vs tocilizumab (4mg/kg)+MTX Placebo (+MTX) vs tocilizumab (8mg/kg)+MTX	-	499	398 (79.8)	52.7 (12.8)	11.7 (9.0)	24	Excluded	SF36	41.1 (11.9)
RA-MOBILITY	Meta-analysis	Placebo (+MTX) vs sarilumab (150mg q2w) Placebo (+MTX) vs sarilumab (200mg q2w)	-	1197	982 (82.0)	50.6 (11.6)	9.1	52	ITT	SF36	38.9 (11.6)
RAPID1	Meta-analysis	Placebo (+MTX) vs certolizumab (400mg) Placebo (+MTX) vs certolizumab (200mg)	2005-2006	982	817 (83.2)	52.0 (11.5)	6.2 (4.3)	12, 24, 52	LOCF	SF36	39.3 (11.2)
RAPID2	Meta-analysis	Placebo (+MTX) vs certolizumab (400mg) Placebo (+MTX) vs	2005-2006	619	505 (81.6)	51.9 (11.6)	6.1 (4.1)	24	LOCF	SF36	39.4 (11.1)

		certolizumab (200mg)									
REFLEX	Meta-analysis	Placebo (+MTX) vs rituximab (2x1000mg)	-	520	420 (80.8)	52.5 (12.4)	11.9 (8.0)	24	LOCF	SF36	39.9 (11.4)
SERENE	Meta-analysis	Placebo (+MTX) vs rituximab (2x500mg)	-	511	418 (81.8)	51.8 (12.7)	7.1 (7.3)	24	Excluded	SF36	41.2 (11.9)
		Placebo (+MTX) vs rituximab (2x1000mg)									
SIRROUND-D	Meta-analysis	Placebo (+MTX) vs sirukumab (50mg q2w or 100mg q4w)	2012-2016	1670	-	-	-	24	ITT	SF36	-
SIRROUND-T	Meta-analysis	Placebo (+MTX) vs sirukumab (50mg q2w or 100mg q4w)	2012-2016	878	712 (81.0)	55.4 (12.2)	12.5 (8.9)	24	ITT	SF36	-
Smolen 2012	Meta-analysis	Placebo (+MTX) vs baricitinib (1mg-8mg)	-	301	-	-	-	12	-	SF36	-
Smolen 2014a	Meta-analysis	Placebo (+MTX) vs sirukumab (100mg eow)	2008-2011	36	25 (69.4)	48.2 (7.0)	7.4 (6.8)	12	LOCF	SF36	37.4 (11.3)
Smolen 2014b	Meta-analysis	Placebo (+MTX) vs sirukumab (25-200mg q4w)	2008-2011	151	133 (88.1)	52.7 (11.3)	10.0 (7.5)	12	LOCF	SF36	37.2 (11.3)
St Clair 2004	Excluded	Placebo (+MTX) vs infliximab (3mg/kg)	2000-2002	1004	713 (71.0)	50.3 (12.7)	0.9 (0.7)	54	LOCF	SF36	-
		Placebo (+MTX) vs infliximab (6mg/kg)									
START	Excluded	Placebo (+MTX) vs infliximab (10mg/kg)	2001-2003	1084	871 (80.4)	52.3	7.5	22	LOCF	SF36	45.1
		Placebo (+MTX) vs infliximab (3mg/kg)									
Strand 2011	Meta-analysis	Placebo (+MTX) vs tofacitinib (5/10mg bid)	-	792	-	-	-	24	Excluded	SF36	41.4 (11.8)
Strand 2012	Narrative Synthesis	Placebo vs secukinumab (25-300mg q4w)	-	237	-	-	-	2, 4, 8, 12, 16	-	SF36	-

Strand 2013	Meta-analysis	Placebo vs decernotinib (25-150mg bid)	-	204	-	56.2 (9.9)	7.7	12	Imputation	SF36	39.0 (11.7)
TACIT	Meta-analysis	cDMARD strategy vs tumour necrosis factor inhibitor strategy	2008-2010	214	144 (67.3)	57.5 (12.0)	Median = 5.2 Range = 1.6-13.4	52	MI	SF36	42.0 (12.0)
TASKi-2	Meta-analysis	Placebo (+MTX) vs fostamatinib (100mg bid) Placebo (+MTX) vs fostamatinib (150mg/day)	-	457	390 (85.3)	52.5 (12.8)	9.2 (8.7)	24	ITT	SF36	40.3 (11.6)
TOWARD	Meta-analysis	Placebo (+MTX) vs tocilizumab (8mg/kg)	2005-2007	1220	997 (81.7)	53.5 (13.0)	9.8 (9.0)	24	ITT	SF36	-
Westhovens 2009	Meta-analysis	Placebo (+MTX) vs abatacept (10mg/kg q4w)	-	509	395 (77.6)	49.9 (12.7)	0.5 (0.6)	52	LOCF	SF36	-

QD once per day. Q4w every 4 weeks. Eow every other week. Wk week. Q8w every 8 weeks. Biw twice a week. Bid twice a day. LOCF Last Observation Carried Forward. ITT Intention to Treat. MTX methotrexate

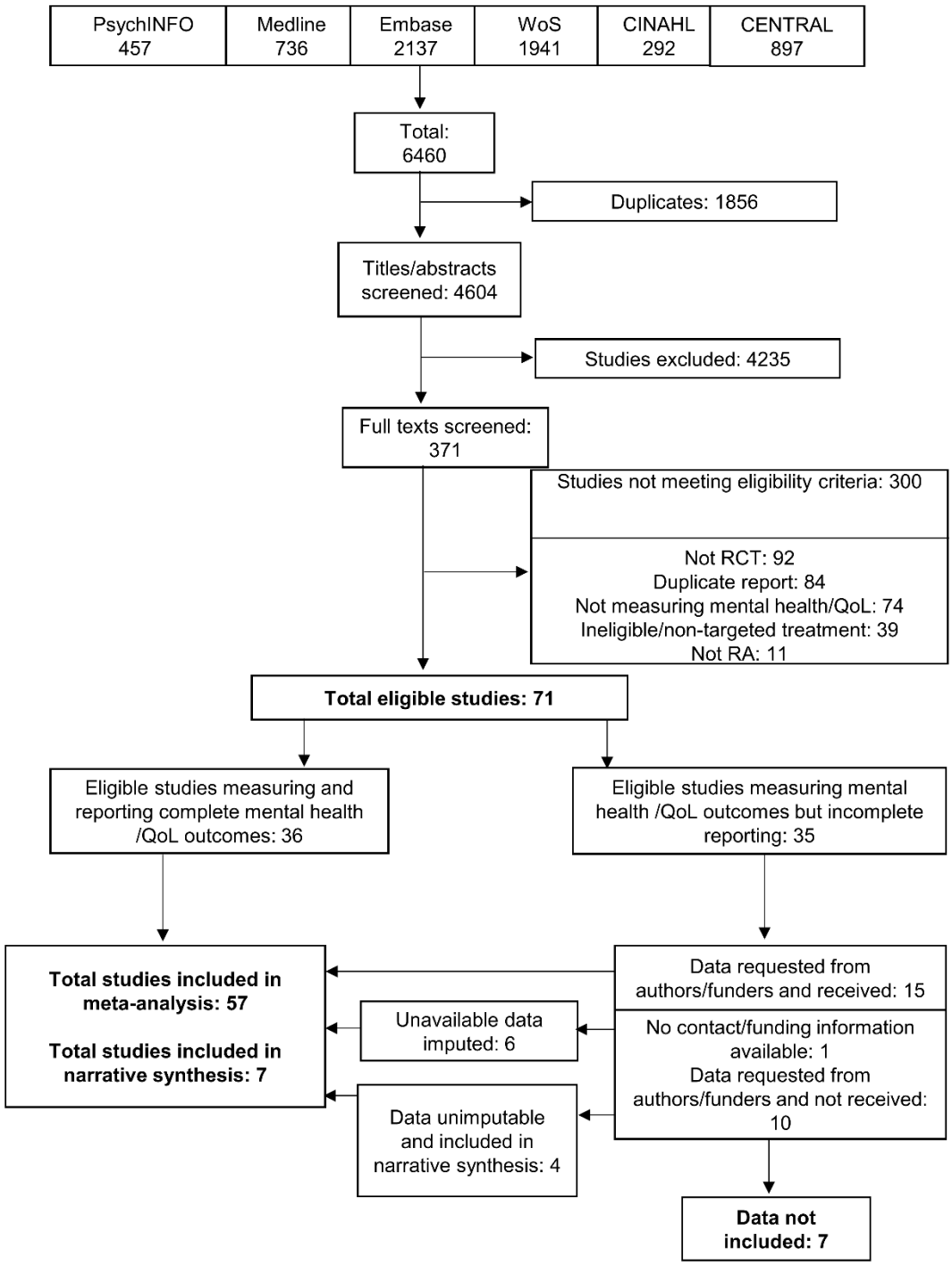

Figure 1. Flow diagram of systematic literature search

Objective 2: The impact of RA treatment on mental health

Results of the PMA, sensitivity and subgroup analyses are shown in table 2. The total analysis involving 57 studies, with no exclusions and all comparators, revealed a statistically significant but modest effect of all treatments on mental HRQoL (MCS) (SMD=0.21). This indicates that, on average, bDMARDs were related to a treatment effect, compared to control treatments, of around one-fifth of a standard deviation, which is equivalent to around a two point difference in MCS units. In comparison, the impact of RA treatments on physical HRQoL (PCS) is somewhat larger (SMD= 0.41) and equivalent to a difference of around four points on the PCS scale. I^2 values reflected moderate-high levels of heterogeneity for both PCS (I^2=76.5) and MCS (I^2=59.2) outcomes. This suggests that estimates may not be robust as an indicator of the population average effect; potentially due to moderating factors, such as differences in trial design.

When limiting the analysis to no-treatment placebo controls, bDMARDs had a substantial benefit for PCS but not MCS outcomes (SMDs = .52 versus .27, respectively). Comparisons with csDMARD controls did not significantly alter the findings from the total analysis (SMDs = .47 versus .24, respectively). For both analyses, heterogeneity levels were reduced compared to the any comparator analysis but remained moderate (>40%) for both MCS and PCS outcomes. Subgroup analysis of unpublished data provided by authors and funders revealed little difference in impact of bDMARDs on MCS and PCS in comparison to background csDMARD control groups compared to all trials.

csDMARDs (typically MTX) was a common comparator against which all bDMARDs had been assessed (see network of comparisons in appendix 5). NMA results for bDMARDs versus csDMARDs are shown in figure 3. These demonstrated consistently small effect sizes for MCS and moderate effects for PCS outcomes. All bDMARDs performed better than csDMARDs for improving MCS and PCS outcomes, although there were no notable differences in outcomes between mode of bDMARD action. Effect sizes for MCS outcomes were typically 50% smaller

than PCS effect sizes. Figure 2 shows the comparator-adjusted funnel plot for the NMA MCS outcome analysis, demonstrating no substantial publication bias.

SUCRA rankings (figure 3) show that for MCS outcomes, out of the drugs considered in the analysis, biologics targeting anti-IL-6 have an 90% probability of being the most effective treatment for MCS outcome; abatacept has an 83% probability of most effectively improving PCS outcomes.

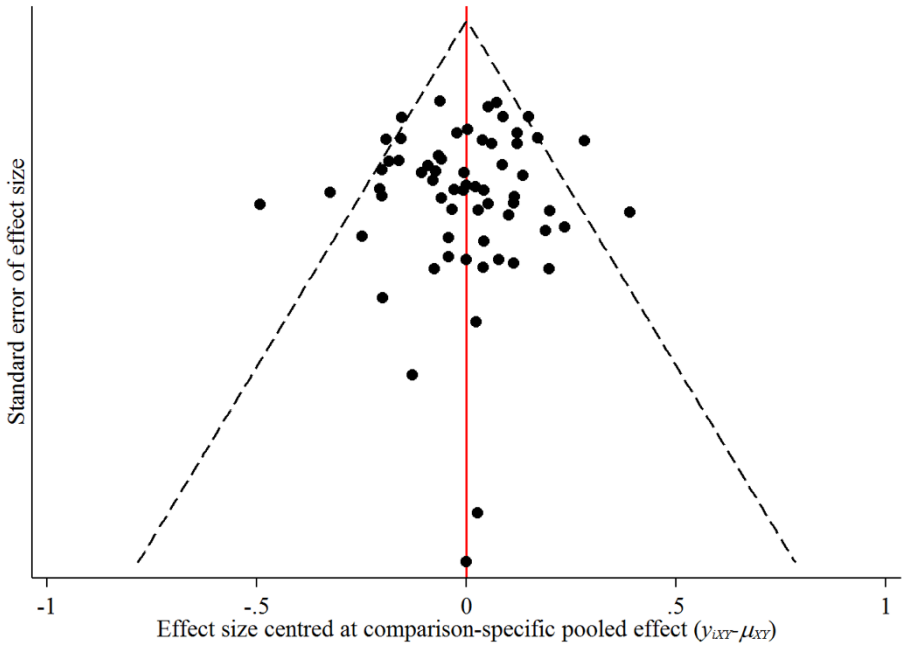

Figure 2. Comparator-adjusted funnel-plot for MCS outcomes

Table 2. Pairwise meta-analysis results with sensitivity and subgroup analysis.

Analysis	Outcome	Number of studies	Number of comparisons	Number of participants	SMD (95% CI)	*p*-value	I^2 statistic (%)
bDMARD vs any comparator*	SF36 MCS	57	67	23,535	0.21 (0.17, 0.25)	<0.001	59.2
	SF36 PCS	55	65	23,108	0.41 (0.35, 0.47)	<0.001	76.5
bDMARD vs no-treatment placebo	SF36 MCS	7	7	2,700	0.27 (0.16, 0.38)	<0.001	41.6
	SF36 PCS	6	6	2,542	0.52 (0.40, 0.64)	<0.001	41.3
bDMARD vs csDMARD	SF36 MCS	44	47	16,678	0.24 (0.19, 0.29)	<0.001	52.9
	SF36 PCS	44	47	16,678	0.47 (0.42, 0.52)	<0.001	57.4
bDMARD vs csDMARD (unpublished)**	SF36 MCS	10	12	3,352	0.22 (0.12, 0.32)	<0.001	46.2
	SF36 PCS	10	12	3,352	0.45 (0.35, 0.55)	<0.001	38.4

bDMARD targeted disease modifying anti-rheumatic drug. csDMARD conventional synthetic disease modifying anti-rheumatic drug. PCS = physical component summary scores (physical quality-of life). MCS = mental component summary scores (mental quality-of life). SMD Standardised Mean Difference. CI = Confidence Intervals. *placebo/csDMARD/steroid/bDMARD **Unpublished data supplied by author/funder.

Objective 3: Variables associated with the impact of RA treatment on mood outcomes.

The results of the meta-regression analyses, including studies to background csDMARD comparators, are provided in Table 3. These results show that sample size, age, proportion female, baseline levels of MH, disease activity, rheumatoid factor (RF) status, year of recruitment and availability of baseline data were not associated with variability in the treatment effect sizes in the PMA results for MCS or PCS outcomes. There was a small but significant positive association between disease duration and MCS outcomes and number of follow-up weeks and PCS outcomes. This indicates that every increased year of disease duration is associated with a 0.04 increase in MCS effect size (i.e. a reduction in treatment efficacy), and every increased week of follow-up time is associated with an increase of 0.01 in PCS effect size.

Risk of bias

The GRADE assessment suggested that the MCS and PCS outcome PMA of bDMARD versus csDMARDs were of moderate quality. Whilst there was no serious indirectness, imprecision or publication bias, few studies were completely without risk of bias and there was moderate heterogeneity. A full summary of the risk of bias assessment is provided in appendix 4.

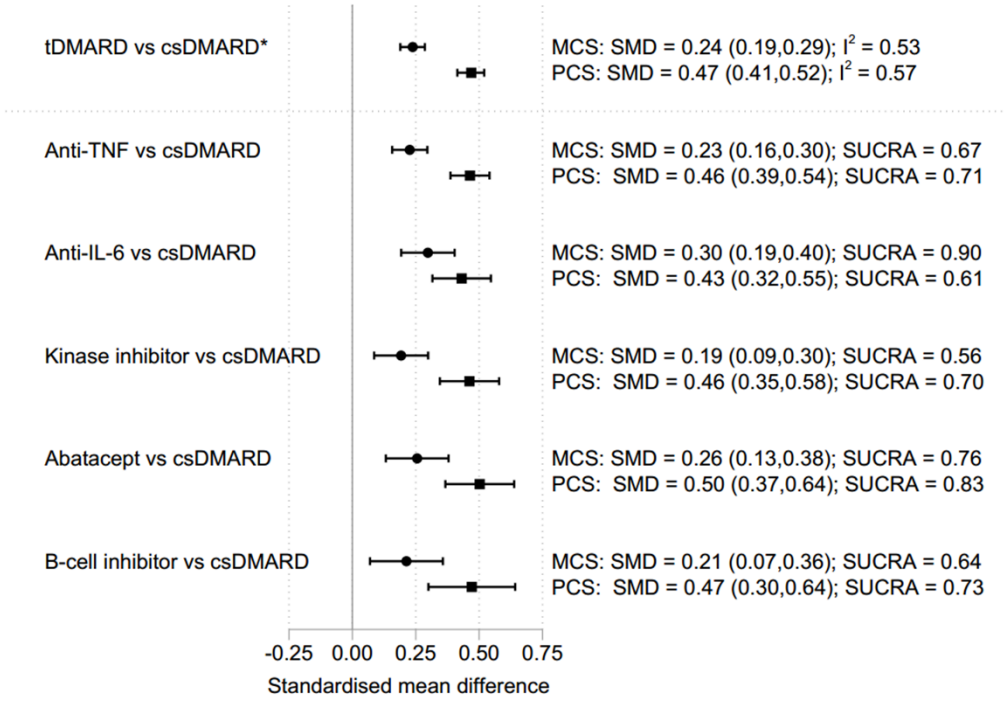

*Figure 3. Estimated pooled treatment effects of biologics therapies on PCS and MCS outcomes. *Total bDAMRD versus csDMARDs pairwise analysis.*

Table 3. Meta-regression of moderators of the impact of RA treatment on MCS and PCS outcomes.

MCS	**Comparison N**	**Study N**	**Participant N**	**SMD (95%CI)**	**p value**	**I2 Statistic (%)**
Comparison analysis: Total versus background DMARD control	44	47	16,678	0.24 (0.19, 0.29)	<0.0001	52.9
Covariates	**Beta**	**SE**	**Lower CI**	**Upper CI**	**p value**	**R-Squared (%)**
Age (continuous)	0.01	0.01	-0.01	0.03	0.244	0.31
Proportion female (continuous)	0.01	0.01	-0.01	0.03	0.683	-0.06
Disease duration (years, continuous)	0.04	0.02	0.00	0.07	0.038	2.83
Early RA (versus established RA)*	-0.28	0.16	-0.61	0.04	0.084	1.68
MCS at baseline (continuous)	-0.03	0.03	-0.08	0.03	0.324	0.04
DAS28 at baseline (continuous)	0.04	0.16	-0.28	0.35	0.806	-0.86
Follow-up time (weeks, continuous)	0.01	0.00	-0.00	0.01	0.056	1.94
Percentage RF positive	0.00	0.01	-0.02	0.02	0.872	-0.89
Year of recruitment start	0.05	0.04	-0.03	0.12	0.213	0.86
Unpublished (versus published)	0.11	0.17	-0.23	0.46	0.519	4.12
PCS	**Comparison N**	**Study N**	**Participant N**	**SMD (95%CI)**	**p value**	**I2 Statistic (%)**
Comparison analysis: Total versus background DMARD control	44	47	16,678	0.47 (0.42, 0.52)	<0.0001	57.4
Covariates	**Beta**	**SE**	**Lower CI**	**Upper CI**	**p value**	**R-Squared (%)**
Age (continuous)	0.01	0.01	-0.01	0.03	0.356	-0.09
Proportion female (continuous)	0.00	0.02	-0.03	0.03	0.986	-0.82
Disease duration (years, continuous)	0.04	0.02	-0.00	0.09	0.066	2.22
Early RA (versus established RA)*	-0.22	0.20	-0.61	0.16	0.250	0.34
MCS at baseline (continuous)	-0.03	0.03	-0.09	0.04	0.373	-0.12
DAS28 at baseline (continuous)	0.06	0.19	-0.31	0.44	0.735	-0.84
Follow-up time (weeks, continuous)	0.01	0.00	-0.00	0.01	0.045	2.27
Percentage RF positive	-0.01	0.01	-0.03	0.02	0.535	-0.73
Year of recruitment start	0.07	0.05	-0.02	0.16	0.129	1.82
Unpublished data (versus published)	0.11	0.22	-0.32	0.54	0.628	6.72

MCS Mental Component Summary. PCS Physical Component Summary. DAS28 28-joint Disease Activity Score. RF Rheumatoid Factor. SMD Standardised Mean Difference. CI Confidence Interval. *Early RA defined as overall study mean disease duration <3 years

DISCUSSION

Despite MH problems being highly prevalent, [3] predictive of worse disease outcomes and treatment response, [8,33] and being highlighted as a priority for outcome measurement by patients, [34,35] 74 (51.0%) of 145 otherwise eligible trials did not measure MH and were excluded from this systematic review. Of the 71 eligible studies indicating that MH had been measured, 35 (49.3%) did not report treatment effect estimates. The results of PMA of 57 trials of targeted treatment show a relatively small but significant impact of bDMARDs on MH assessed by the SF36. The impact of targeted RA treatment on SF36-MCS was approximately half the effect seen in SF36-PCS. The largest effect size for MCS outcomes was 0.30, found for the anti-IL-6 versus csDMARD comparison; the lowest effect size was 0.19, found in the Kinase inhibitors.

To date, TNFα has been the primary focus of research investigating the inflammatory mechanisms involved in the presence of depressive symptomatology. Infliximab has been recently investigated as an anti-depressant in treatment-resistant depression, [14] and the impact of anti-TNF medications on depression outcomes in chronic physical conditions has been addressed in a recent systematic review and meta-analysis of six trials. [15] Building on this review, [15] we focused only on RA, but included broader conceptualisations of MH and more treatment types. By including treatments with varied modes of action, we hoped to pinpoint the mechanism through which RA treatment may have benefits for MH. However, we failed to find any major variations between treatment modes of action. Whilst we found one of the largest effects on MH for treatments targeting IL-6, the smallest effect size was observed for anti-TNF treatments. Therefore, it remains largely unclear as to the extent to which improvements in MH are through bDMARDs directly impacting inflammatory pathways, or simply indirectly through the reduction in pain and disability.

Meta-regression analysis identified a small but significant association between disease duration and MCS effect size, and the largest (although non-significant) R-squared value for comparing data which had been published online versus unpublished data which was

requested from authors. Although we found no clear evidence of publication bias in our funnel plots, there may be a tendency for non-significant mental health outcomes to be omitted from published papers. [36]

This review used reproducible and rigorous methods to collate and synthesize the data in this field. We included many trials, representing >20,000 patients, and study quality was relatively high. There are some restrictions which limit the interpretation of our results. We used broad inclusion criteria for the entry of studies into this review, preferring to use sensitivity and subgroup analyses and meta-regression to examine sources of statistical heterogeneity in the PMA, which was substantial. In addition to heterogeneity due to the different types of bDMARDs included, heterogeneity may also be explained by the comparator used, plus variability in disease duration and length of follow-up between studies. Another, source of heterogeneity may be that we did not restrict our focus to trials specifically recruiting patients with low mood at baseline. The overall mean MCS score at baseline was 42.2, with 20.8% of studies reporting a mean MCS score reaching below a threshold of 40, indicated as a threshold for possible mood disorder. [37] Most patients included in the studies may not have had mood disorder at baseline, restricting potential to find an 'anti-depressant' effect.

NMA methodologies are being more widely used in medical research, however there are limitations to the technique which need addressing. Firstly, it is important to highlight that, as treatment allocations have been randomised within (not between) trials, NMA can only provide observational evidence [38]. NMA assumes transitivity (whether any patient could be given any treatment in the network) and consistency (similar estimates obtained from direct and indirect comparisons). Our focus on bDMARDS, which are relatively recently developed, typically involve similar inclusion criteria, and generally are considered to be equally efficacious [39], limits the potential for violation of the transitivity assumption. Regarding the consistency assumption, examining loop specific heterogeneity we found no specific cause for concern.

Despite not limiting our search strategy to the SF36, we identified the SF36 as the most commonly-used tool for measuring mental health, and data from this were prioritised to allow meaningful comparison across studies. Whilst this measure allows interesting comparison between mental and physical QoL outcomes, it is important to acknowledge that the SF36 MCS captures a broader conceptualisation of mental health-related quality of life. This includes symptoms of depression and anxiety but also vitality/fatigue and impacts on social and emotional functioning. [24] Future research may benefit from identifying subgroups of patients who may be susceptible to experiencing MH benefits following RA treatments and understanding how these patients may differ from those who are more resistant to improvement. This may provide useful clinical information to anticipate treatment response, as improvement in MH in turn is likely to further impact physical symptom experiences [33]. This approach may also identify potentially useful intervention targets. A focus on RA patients with symptoms of MH at baseline may provide insight into any benefits of RA treatment on a subgroup of people with both heightened inflammation and psychological disorder.

Conclusion

Advances in RA treatment have resulted in significant improvements in specific outcomes: the delay of radiographic damage and reduction of inflammation and adverse events. [40] However this review demonstrates that relying on RA pharmacotherapy alone may not meaningfully improve MH outcomes. MH is treatable in patients with physical illness, [41,42] and the measurement and management of MH throughout the course of treatment as part of routine practice is recommended. [43] Our results suggest that MH in patients with RA must be addressed and are unlikely to resolve with effective RA pharmacological disease management alone. Providing integrated, dedicated MH care within routine practice is essential to achieve parity of esteem, valuing mental and physical health equally.

CONTRIBUTORS

FM was responsible for the conception and design of the review as well as the overall conduct and publication write-up. JG provided rheumatological expertise where required and aided in

the interpretation and understanding of results. MH supervised the conduct of the review, providing advice and assistance in the conceptualisation and interpretation of results. ER performed all secondary independent study screening and data extraction and advised on the conduct of the review. ICS and SS contributed to the conception and design of the review, as well as assisting with rheumatological advice and interpretation of findings. SN performed all analyses and contributed to the design, development and conduct of the review including interpretation of results. All authors reviewed the final submitted version of the manuscript.

ACKNOWLEDGEMENTS

This paper represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

REFERENCES

1 Kvien TK. Epidemiology and burden of illness of rheumatoid arthritis. *Pharmacoeconomics* 2004;**22**:1–12.

2 Matcham F, Scott IC, Rayner L, *et al.* The impact of rheumatoid arthritis on quality-of-life assessed using the SF-36: A systematic review and meta-analysis. *Semin Arthritis Rheum* 2014;**44**:123–30. doi:10.1016/j.semarthrit.2014.05.001

3 Matcham F, Rayner L, Steer S, *et al.* The prevalence of depression in rheumatoid arthritis: a systematic review and meta-analysis. *Rheumatology (Oxford)* 2013;**52**:2136–48. doi:10.1093/rheumatology/ket169

4 Rayner L, Matcham F, Hutton J, *et al.* Embedding integrated mental health assessment and management in general hospital settings: feasibility, acceptability and the prevalence of common mental disorder. *Gen Hosp Psychiatry* 2014;**36**:318–24. doi:10.1016/j.genhosppsych.2013.12.004

5 Waraich P, Goldner E, Somers J, *et al.* Prevalence and incidence studies of mood disorders: A systematic review of the literature. *Can J Psychiatry* 2009;**49**:124–38.

6 van den Hoek J, Boshuizen HC, Roorda LD, *et al.* Somatic comorbidities and comorbid depression are associated with mortality in patients with rheumatoid arthritis: A 14-year prospective cohort study. *Arthritis Care Res (Hoboken)* 2015;**68**:1055–60. doi:10.1002/ACR.22812

7 Löwe B, Willand L, Eich W, *et al.* Psychiatric comorbidity and work disability in patients with inflammatory rheumatic diseases. *Psychosom Med* 2004;**66**:395–402.

8 Matcham F, Norton S, Scott DL, *et al.* Symptoms of depression and anxiety predict treatment response and long-term physical health outcomes in rheumatoid arthritis: secondary analysis of a randomized controlled trial. *Rheumatology (Oxford)* 2016;**55**:268–78. doi:10.1093/rheumatology/kev306

9 Matcham F, Ali S, Irving K, *et al.* Are depression and anxiety associated with disease activity in rheumatoid arthritis? A prospective study. *BMC Musculoskelet Disord* 2016;**17**. doi:10.1186/s12891-016-1011-1

10 Karpouzas GA, Dolatabadi S, Moran R, *et al.* Correlates and predictors of disability in vulnerable US Hispanics with rheumatoid arthritis. *Arthritis Care Res (Hoboken)* 2012;**64**:1274–81. doi:10.1002/acr.21689

11 Kojima M, Kojima T, Suzuki S, *et al.* Depression, inflammation, and pain in patients with rheumatoid arthritis. *Arthritis Rheum* 2009;**61**:1018–24. doi:10.1002/art.24647

12 Matcham, Ali S, Hotopf M, *et al.* Psychological correlates of fatigue in rheumatoid arthritis: A systematic review. *Clin Psychol Rev* 2015;**39**:16–29. doi:10.1016/j.cpr.2015.03.004

13 Liu Y, Ho RC-M, Mak A. Interleukin (IL)-6, tumour necrosis factor alpha (TNF-α) and soluble interleukin-2 receptors (sIL-2R) are elevated in patients with major depressive disorder: a meta-analysis and meta-regression. *J Affect Disord* 2012;**139**:230–9. doi:10.1016/j.jad.2011.08.003

14 Raison CL, Rutherford RE, Woolwine BJ, *et al.* A randomized controlled trial of the tumor necrosis factor antagonist infliximab for treatment-resistant depression: the role of baseline inflammatory biomarkers. *JAMA psychiatry* 2013;**70**:31–41. doi:10.1001/2013.jamapsychiatry.4

15 Abbott R, Whear R, Nikolaou V, *et al.* Tumour necrosis factor-α inhibitor therapy in chronic physical illness: A systematic review and meta-analysis of the effect on depression and anxiety. *J Psychosom Res* 2015;**79**:175–84. doi:10.1016/j.jpsychores.2015.04.008

16 Diffin JG, Lunt M, Marshall T, *et al.* Has the severity of rheumatoid arthritis at presentation diminished over time? *J Rheumatol* 2014;**41**:1590–9. doi:10.3899/jrheum.131136

17 Smolen JS, Breedveld FC, Burmester GR, *et al.* Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. *Ann Rheum Dis* 2016;**75**:3–15. doi:10.1136/annrheumdis-2015-207524

18 Smolen JS, Aletaha D, Bijlsma JWJ, *et al.* Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;**69**:631–7. doi:10.1136/ard.2009.123919

19 National Institute for Health and Care Excellence. Rheumatoid arthritis: The management of rheumatoid arthritis in adults. Clinical guideline (CG79). http://www.nice.org.uk/guidance/CG79: : NICE 2015. http://www.nice.org.uk/guidance/CG79 (accessed 20 Aug 2014).

20 Kekow J, Moots RJ, Emery P, *et al.* Patient-reported outcomes improve with etanercept plus methotrexate in active early rheumatoid arthritis and the improvement is strongly associated with remission: the COMET trial. *Ann Rheum Dis* 2010;**69**:222–5. doi:10.1136/ard.2008.102509

21 Scott IC, Ibrahim F, Lewis CM, *et al.* Impact of intensive treatment and remission on health-related quality of life in early and established rheumatoid arthritis. *RMD open* 2016;**2**:e000270. doi:10.1136/rmdopen-2016-000270

22 Sanderson T, Morris M, Calnan M, *et al.* Patient perspective of measuring treatment efficacy: the rheumatoid arthritis patient priorities for pharmacologic interventions outcomes. *Arthritis Care Res (Hoboken)* 2010;**62**:647–56. doi:10.1002/acr.20151

23 Kalyoncu U, Dougados M, Daurès J-P, *et al.* Reporting of patient-reported outcomes in recent trials in rheumatoid arthritis: a systematic literature review. *Ann Rheum Dis* 2009;**68**:183–90. doi:10.1136/ard.2007.084848

24 Ware JEJ, Snow K, Kosinski M. *SF-36 Health Survey: Manual and interpretation guide.* Boston: : Health Institute, New England Medical Centre. 1993.

25 Moher D. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Ann Intern Med* 2009;**151**:264. doi:10.7326/0003-4819-151-4-200908180-00135

26 Ware JE, Kosinski M, Keller S. *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston, MA: : The Health Institute 1994.

27 Taft C, Karlsson J, Sullivan M. Do SF-36 summary component scores accurately summarize subscale scores? *Qual Life Res* 2001;**10**:395–404. doi:10.1023/A:1012552211996

28 Higgins JPT, Altman DG, Gøtzsche PC, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Br Med J* 2011;**343**:d5928. doi:10.1136/bmj.d5928

29 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: : Erlbaum 1988.

30 Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. Cochrane Collab. 2011;:www.cochrane-handbook.org.

31 Higgins JPT, Thompson SG, Deeks JJ, *et al.* Measuring inconsistency in meta-analyses. *Br Med J* 2003;**327**:557–60. doi:10.1136/bmj.327.7414.557

32 Mavridis D, Giannatsi M, Cipriani A, *et al.* A primer on network meta-analysis with emphasis on mental health. *Evid Based Ment Heal* 2015;**18**:40–6. doi:10.1136/EB-2015-102088

33 Matcham F, Davies R, Hotopf M, *et al.* The relationship between depression and biologic treatment response in rheumatoid arthritis: An analysis of the British Society for Rheumatology Biologics Register. *Rheumatology* Published Online First: 13 February 2018. doi:10.1093/rheumatology/kex528

34 Kirwan JR, Hewlett SE, Heiberg T, *et al.* Incorporating the patient perspective into outcome assessment in rheumatoid arthritis--progress at OMERACT 7. *J Rheumatol* 2005;**32**:2250–6.http://www.ncbi.nlm.nih.gov/pubmed/16265712 (accessed 27 Jul 2016).

35 Kirwan J, Heiberg T, Hewlett S, *et al.* Outcomes from the Patient Perspective Workshop at OMERACT 6. *J Rheumatol* 2003;**30**:868–72.http://www.ncbi.nlm.nih.gov/pubmed/12672218 (accessed 27 Jul 2016).

36 Smyth RMD, Kirkham JJ, Jacoby A, *et al.* Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. *Br Med J* 2011;**342**:c7153. doi:10.1136/bmj.c7153

37 Matcham F, Norton S, Steer S, *et al.* Usefulness of the SF-36 Health Survey in screening for depressive and anxiety disorders in rheumatoid arthritis. *BMC Musculoskelet Disord* 2016;**17**:224. doi:10.1186/s12891-016-1083-y

38 Cipriani A, Higgins JPT, Geddes JR, *et al.* Conceptual and Technical Challenges in Network Meta-analysis. *Ann Intern Med* 2013;**159**:130. doi:10.7326/0003-4819-159-2-201307160-00008

39 Stevenson M, Archer R, Tosh J, *et al.* Adalimumab, etanercept, infliximab, certolizumab pegol, golimumab, tocilizumab and abatacept for the treatment of rheumatoid arthritis not previously treated with disease-modifying antirheumatic drugs and after the failure of conventional disease-modifyin. *Health Technol Assess (Rockv)* 2016;**20**:1–610.

40 Hazlewood GS, Barnabe C, Tomlinson G, *et al.* Methotrexate monotherapy and methotrexate combination therapy with traditional and biologic disease modifying antirheumatic drugs for rheumatoid arthritis: abridged Cochrane systematic review and network meta-analysis. *Br Med J* 2016;**353**:i1777. doi:10.1136/bmj.i1777

41 Matcham F, Rayner L, Hutton J, *et al.* Self-help interventions for symptoms of depression, anxiety and psychological distress in patients with physical illnesses: A systematic review and meta-analysis. Clin. Psychol. Rev. 2014;**34**:141–57. doi:10.1016/j.cpr.2014.01.005

42 Rayner L, Price A, Evans A, *et al.* Antidepressants for depression in physically ill people. Cochrane Database Syst. Rev. 2010;:CD007503. doi:10.1002/14651858.CD007503.pub2.

43 National Institute for Health and Care Excellence. Depression in adults with a chronic physical health problem: recognition and management (clinical guideline CG91). http://www.nice.org.uk/guidance/cg91: : NICE 2009. http://www.nice.org.uk/guidance/cg91 (accessed 20 Aug 2014).